

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 24, 2004, 22:29:26 ; Search time 22 Seconds  
(without alignments)

734.497 Million cell updates/sec

Title: US-09-900-038a-1

Perfect score: 1590  
Sequence: 1 MNYSIINSYNEPLNYRDS.....LINDINILVLTKEGKQSD 313

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: /cgm2\_6/ptodata/2/1aa/5A\_COMB.pep.\*  
2: /cgm2\_6/ptodata/2/1aa/5B\_COMB.pep.\*  
3: /cgm2\_6/ptodata/2/1aa/6A\_COMB.pep.\*  
4: /cgm2\_6/ptodata/2/1aa/6B\_COMB.pep.\*  
5: /cgm2\_6/ptodata/2/1aa/6C\_COMB.pep.\*  
6: /cgm2\_6/ptodata/2/1aa/6D\_COMB.pep.\*  
7: /cgm2\_6/ptodata/2/1aa/6E\_COMB.pep.\*  
8: /cgm2\_6/ptodata/2/1aa/6F\_COMB.pep.\*  
9: /cgm2\_6/ptodata/2/1aa/6G\_COMB.pep.\*  
10: /cgm2\_6/ptodata/2/1aa/6H\_COMB.pep.\*  
11: /cgm2\_6/ptodata/2/1aa/6I\_COMB.pep.\*  
12: /cgm2\_6/ptodata/2/1aa/6J\_COMB.pep.\*  
13: /cgm2\_6/ptodata/2/1aa/6K\_COMB.pep.\*  
14: /cgm2\_6/ptodata/2/1aa/6L\_COMB.pep.\*  
15: /cgm2\_6/ptodata/2/1aa/6M\_COMB.pep.\*  
16: /cgm2\_6/ptodata/2/1aa/6N\_COMB.pep.\*  
17: /cgm2\_6/ptodata/2/1aa/6O\_COMB.pep.\*  
18: /cgm2\_6/ptodata/2/1aa/6P\_COMB.pep.\*  
19: /cgm2\_6/ptodata/2/1aa/6Q\_COMB.pep.\*  
20: /cgm2\_6/ptodata/2/1aa/6R\_COMB.pep.\*  
21: /cgm2\_6/ptodata/2/1aa/6S\_COMB.pep.\*  
22: /cgm2\_6/ptodata/2/1aa/6T\_COMB.pep.\*  
23: /cgm2\_6/ptodata/2/1aa/6U\_COMB.pep.\*  
24: /cgm2\_6/ptodata/2/1aa/6V\_COMB.pep.\*  
25: /cgm2\_6/ptodata/2/1aa/6W\_COMB.pep.\*  
26: /cgm2\_6/ptodata/2/1aa/6X\_COMB.pep.\*  
27: /cgm2\_6/ptodata/2/1aa/6Y\_COMB.pep.\*  
28: /cgm2\_6/ptodata/2/1aa/6Z\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	234	14.7	301	US-09-495-406-15	Sequence 15, Appl
2	216	13.6	534	US-09-134-000C-5087	Sequence 5087, Ap
3	214.5	13.5	1056	US-09-134-000C-5086	Sequence 5086, Ap
4	210.5	13.2	348	US-08-312-387B-3	Sequence 3, Appl
5	210.5	13.2	348	US-08-312-387B-11	Sequence 11, Appl
6	210.5	13.2	348	US-08-683-426-3	Sequence 3, Appl
7	210.5	13.2	348	US-08-683-426-11	Sequence 11, Appl
8	210.5	13.2	348	US-08-683-458-3	Sequence 3, Appl
9	210.5	13.2	348	US-08-683-458-11	Sequence 11, Appl
10	210.5	13.2	348	US-08-878-360-3	Sequence 3, Appl
11	210.5	13.2	348	US-08-878-360-11	Sequence 11, Appl
12	210.5	13.2	348	US-08-478-140B-8	Sequence 8, Appl
13	210.5	13.2	348	US-08-478-140B-3	Sequence 3, Appl
14	210.5	13.2	348	US-09-333-412-3	Sequence 3, Appl
15	210.5	13.2	348	US-09-333-412-11	Sequence 11, Appl
16	210.5	13.2	348	US-09-338-943-3	Sequence 3, Appl
17	210.5	13.2	348	US-09-338-943-8	Sequence 8, Appl
18	208	13.1	301	US-09-252-991A-33096	Sequence 33096, A
19	200	12.6	416	US-09-495-406-25	Sequence 25, Appl
20	199	12.5	416	US-09-489-039A-8195	Sequence 8195, Ap
21	190.5	12.0	333	US-09-107-532A-5123	Sequence 5123, Ap
22	187.5	11.8	277	US-09-107-532A-7213	Sequence 7213, Ap
23	186.5	11.7	327	US-09-107-532A-6181	Sequence 6181, Ap
24	184	11.6	303	US-09-495-406-17	Sequence 17, Appl
25	176.5	11.1	337	US-08-312-387B-5	Sequence 5, Appl
26	176.5	11.1	337	US-08-312-387B-12	Sequence 12, Appl
27	176.5	11.1	337	US-08-683-426-5	Sequence 5, Appl

28	176.5	11.1	337	1	US-08-683-426-12	Sequence 12, Appl
29	176.5	11.1	337	1	US-08-683-458-5	Sequence 5, Appl
30	176.5	11.1	337	1	US-08-683-458-12	Sequence 12, Appl
31	176.5	11.1	337	2	US-08-878-360-5	Sequence 5, Appl
32	176.5	11.1	337	2	US-08-878-360-12	Sequence 12, Appl
33	176.5	11.1	337	3	US-08-478-140B-5	Sequence 5, Appl
34	176.5	11.1	337	3	US-09-333-412-5	Sequence 5, Appl
35	176.5	11.1	337	4	US-09-333-412-12	Sequence 12, Appl
36	176.5	11.1	337	4	US-09-338-943-5	Sequence 5, Appl
37	171.5	10.8	721	4	US-09-107-532A-6889	Sequence 6889, Ap
38	168	10.6	702	4	US-09-437-277-1	Sequence 1, Appl
39	165	10.4	248	4	US-09-107-532A-4568	Sequence 4568, Ap
40	165	10.4	965	4	US-09-437-277-3	Sequence 3, Appl
41	161	10.1	324	1	US-08-597-226-10	Sequence 10, Appl
42	161	10.1	324	1	US-08-746-682A-10	Sequence 10, Appl
43	160.5	10.1	303	4	US-09-252-991A-29155	Sequence 29155, A
44	160	10.1	281	3	US-08-961-083-196	Sequence 196, App
45	160	10.1	281	4	US-09-536-784-196	Sequence 196, App

## ALIGNMENTS

RESULT 1	US-09-495-406-15	Application US/09495406
Sequence 15, Appl	US-09-495-406-15	Sequence 15, Appl
Patent No. 6503744		
GENERAL INFORMATION:		
APPLICANT: Gilbert, Michel		
APPLICANT: Wakarchuk, Warren W.		
TITLE OF INVENTION: National Research Council of Canada		
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of		
FILE REFERENCE: 019633-000110US		
CURRENT APPLICATION NUMBER: US/09/495,406		
CURRENT FILING DATE: 2000-01-31		
PRIOR APPLICATION NUMBER: US 60/118,213		
PRIOR FILING DATE: 1999-02-01		
NUMBER OF SEQ ID NOS: 35		
SOFTWARE: Patentn Ver. 2.1		
SEQ ID NO 15		
LENGTH: 301		
TYPE: PRT		
ORGANISM: Campylobacter jejuni		
US-09-495-406-15		
Query Match	14.7%	Score 234; DB 4; Length 301;
Best Local Similarity	28.2%	Pred. No. 1.3e-13;
Matches	88; Conservative	53; Mismatches 135; Indels 36; Gaps 13;
4	SIIMSYNEPLNYRDSVESINQTLTDFEFITVDNPSRGLKQFLREYSVVDNRKLT	63
5	SILPITYNVE-QYIARAISICINQTFKQIE-IIVDDCGDINSINIAKESKDKRIKII	62
64	LINEENGLASSLUNKVKSKEVIFPMDDDISYSPRDKOIRFMEBNSLDSATLIEI	123
63	HNEKNGLRANRYEGVKNVSPYIMFDDDLLEINACECKIKIDBDQ-EVDLYFENAI	121
124	DQKSNLYVQKRSNK-IYVINDIRKMLNRSIIAHPW-CYKCVFEDKMGYRDLVPE	180
122	VEENSVSYKPFNSGFYSKEFEVKIIRKQMLYTMWKILRKLYLEA-----	171
181	DYDFARGLADPFKIGLNLKLLQYRLNNGSQQNKFKQYYSALIOFYEKSYIDT	240
172	FAELRIEKVKIKIMADVLLYPM---ISQAKIA-YMCKVFLHYVNNNSICNT	222
241	K---ITNYFOEVIKRYTQOE--LSKYFEIKSPSITIR-KYICLYLYKSPVLRRL	293
223	KNEVLKNNIQELQVLNLYLRQNYILNKYC---SVLYVILKILVLYQYKIRKRLMVL	279
294	LINDINILVLT 305	
280	LAK-INILTLKI 290	

RESULT 2  
US-09-134-000C-5087  
Sequence 5087, Application US/09134000C  
Patent No. 6617156  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 032796-0032  
CURRENT APPLICATION NUMBER: US/09/134,000C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/055,778  
PRIOR FILING DATE: 1997-08-15  
NUMBER OF SEQ ID NOS: 6812  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 5087  
LENGTH: 534  
TYPE: PRT  
ORGANISM: Enterococcus faecalis  
US-09-134-000C-5087

Query Match 13.6%; Score 216; DB 4; Length 534;  
Best Local Similarity 23.3%; Pred. No. 1.3e-11;  
Matches 79; Conservative 64; Mismatches 138; Indels 58; Gaps 10;

QY 4 SIMSVNPELNYVDVSESLNQTLDPEFIIVDNPSSGDLKQFLTEYSVDNRKIL 63  
DB 196 SIAMPVYVEEKWRLCLDSTLNVNWEICMADDSIDPNVKILTEYQOLDERIRV 255  
QY 64 LNEENIGLASSLNKAVKISKGEYIFRMDADDISYFRDQIRPWE-ENSLDFSATLIEL 122  
DB 256 FRENGHISEATNSALSTATGPIFMDNDDELAPQALYEVVKALNTPDPIYTDSDX 315  
QY 123 IDQGNLYVKQESKQITLNDIRKMLNRSILAHPTVCYKKYFDKLMGRDLY-PEVD 181  
DB 316 IDMDGN-----RSDPAFKPMSPDLLIGTVYISH-LGVYRSILIEBIGFRKGYESQD 368  
QY 182 YDPAIRGALADPK--IGILNKVLLQVRLNENGISQTNKFKQYISA----- 225  
DB 369 YDLVRFTEKTKRIKIPKLVYKRLPSTAVDQSGKGYAFEGALRAVDALVRGI 428  
QY 226 -----ILDDPYEK-----SYDITKITNYFOGYIKKQYQELS 261  
DB 429 NGATHGANGLYVYVDISEKLVSIILPTKNGYKDVQRCVSIIE---KTYQNYEI- 484  
QY 262 KYFELKSPSITIRKLYICLVYFKSPVLVRLINDINI 300  
DB 485 ----IMADNGSTDPRKAE-LYAEFEQQLPGRFFVESIDI 518

RESULT 3  
US-09-134-000C-5086  
Sequence 5086, Application US/09134000C  
Patent No. 6617156  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 032796-0032  
CURRENT APPLICATION NUMBER: US/09/134,000C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/055,778  
PRIOR FILING DATE: 1997-08-15  
NUMBER OF SEQ ID NOS: 6812  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 5086  
LENGTH: 1056  
TYPE: PRT  
ORGANISM: Enterococcus faecalis  
US-09-134-000C-5086

Query Match 13.5%; Score 214.5; DB 4; Length 1056;  
Best Local Similarity 23.1%; Pred. No. 4.8e-11;  
Matches 69; Conservative 58; Mismatches 131; Indels 41; Gaps 6;

QY 4 SIMSVNPELNYVDVSESLNQTLDPEFIIVDNPSSGDLKQFLTEYSVDNRKIL 63  
DB 566 SVAPVYVEEKWRLCLDSTLNVNWEICMADDSIDPNVKILTEYQOLDERIRV 625  
QY 64 LNEENIGLASSLNKAVKISKGEYIFRMDADDISYFRDQIRPWE-ENSLDFSATLIEL 122  
DB 626 FRENGHISEATNSALSTATGPIFMDNDDELAPQALYEVVKALNTPDPIYTDSDX 685  
QY 123 IDQGNLYVKQESKQITLNDIRKMLNRSILAHPTVCYKKYFDKLMGRDLY-PEVD 181  
DB 686 ITENG-----RRFAFKSDPNPELILNHYITHFV-VVRDILEKVGINSAYNAQD 738  
QY 182 YDPAIRGALADPKILNKVLLQVRLNENGISQTNKFKQYISA-----ILDDPYEKSYID 238  
DB 739 YDFVLRATEQATKIRIGIPGMVYHWALESSTALNPESKGYAVVAGQKAVQATERGLKA 798  
QY 239 ITKITNYQOEYIKRYQQLSKFELKSTPITIRKLYICLVYFKSPVLVRLIND 297  
DB 799 QVEIAEFYGSYKIN-----LYLXHPVWSLITND 828

RESULT 4  
US-08-312-387B-3  
Sequence 3, Application US/08312387B  
Patent No. 5545353

GENERAL INFORMATION:  
APPLICANT: Gotschlich, Emil C.  
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/312,387B  
FILING DATE: July 7, 1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-095  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-312-387B-3

Query Match 13.2%; Score 210.5; DB 1; Length 348;  
Best Local Similarity 27.8%; Pred. No. 2.3e-11;  
Matches 86; Conservative 55; Mismatches 123; Indels 45; Gaps 15;

QY 4 SIMSVNPELNYVDVSESLNQTLDPEFIIVDNPSSGDLKQFLTEYSVDNRKIL 63  
DB 6 SVLICAYNVE-KYFAQSILAAVNVQTNWINDILYVDDSGTDL-ALAKQFORDSRIKIL 63

QY 64 LNEENIGGLASSINKAV-KISK-----GEYIFRMDADDISYSPRFDKQIRFMEENS--LDF 115  
DB 64 AQAONSGILPSLNLGLDELANSGGGGGYIARTDADIASBGMTEKIVGENEKORSTIAM 123  
QY 116 SATLIEIDOK-GNLVYKORESNKIYLT---NDIRKMLNRSILAHPTWCVKKYPDKL 170  
DB 124 GAMLEVLSEKQGNLRAHHGKIKWKPTRHEDIAAFPPGPNPHNNTIMRERSVIDGG 183  
QY 171 MGY---RDVPEVDYDF---AIRGALADFKIGLNLKVLQYRLNENGISQTKNFKQYIY 223  
DB 184 LRYDTERMW--AEDYQFWYDVSKLGRLAYYP---EALVKYRLHANQVSSKHSVRQH-- 234  
QY 224 SAILDIFYEKSYIDITKITNYFOEYVKKRYT---QOELSKYFEL--KSTPSITIRKL 277  
DB 235 -EIAQIOCK-----TARNDFLOSMGFRTFDSLEVRQTKAAAYELPEKDLPEEDPERA 286  
QY 278 YICLYLYFK 286  
DB 287 RRFVQCFK 295

## RESULT 5

US-08-312-387B-11  
; Sequence 11, Application US/08312387B  
; Patent No. 5545553  
; GENERAL INFORMATION:  
; APPLICANT: Gotschlich, Emil C.  
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF  
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/312,387B  
; FILING DATE: July 7, 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-095  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 348 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-312-387B-11

Query Match 13.2%; Score 210.5; DB 1; Length 348;  
Best Local Similarity 27.8%; Pred. No. 2.3e-11;  
Matches 86; Conservative 55; Mismatches 123; Indels 45; Gaps 15;  
QY 4 SIIMSYYNEPLNYRDSVESILNQLTDFEFIVINPBRGLKQFLTESVVDNRKIKL 63  
DB 6 SYVICAYNVE-KYFAQSLAAVNVQTRNLDILIVDGSIDGTL-AIAKDFQKDSRIKIL 63  
QY 64 LNEENIGGLASSINKAV-KISK-----GEYIFRMDADDISYSPRFDKQIRFMEENS--LDF 115

DB 64 AQAONSGILPSLNLGLDELANSGGGGGYIARTDADIASBGMTEKIVGENEKORSTIAM 123  
QY 116 SATLIEIDOK-GNLVYKORESNKIYLT---NDIRKMLNRSILAHPTWCVKKYPDKL 170  
DB 124 GAMLEVLSEKQGNLRAHHGKIKWKPTRHEDIAAFPPGPNPHNNTIMRERSVIDGG 183  
QY 171 MGY---RDVPEVDYDF---AIRGALADFKIGLNLKVLQYRLNENGISQTKNFKQYIY 223  
DB 184 LRYDTERMW--AEDYQFWYDVSKLGRLAYYP---EALVKYRLHANQVSSKHSVRQH-- 234  
QY 224 SAILDIFYEKSYIDITKITNYFOEYVKKRYT---QOELSKYFEL--KSTPSITIRKL 277  
DB 235 -EIAQIOCK-----TARNDFLOSMGFRTFDSLEVRQTKAAAYELPEKDLPEEDPERA 286  
QY 278 YICLYLYFK 286  
DB 287 RRFVQCFK 295

## RESULT 6

US-08-683-426-3  
; Sequence 3, Application US/08683426  
; Patent No. 5705367  
; GENERAL INFORMATION:  
; APPLICANT: Gotschlich, Emil C.  
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF  
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/683,426  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/312,387  
; FILING DATE: September 26, 1994  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-095B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 348 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-683-426-3

Query Match 13.2%; Score 210.5; DB 1; Length 348;  
Best Local Similarity 27.8%; Pred. No. 2.3e-11;  
Matches 86; Conservative 55; Mismatches 123; Indels 45; Gaps 15;  
QY 4 SIIMSYYNEPLNYRDSVESILNQLTDFEFIVINPBRGLKQFLTESVVDNRKIKL 63  
DB 6 SYVICAYNVE-KYFAQSLAAVNVQTRNLDILIVDGSIDGTL-AIAKDFQKDSRIKIL 63  
QY 64 LNEENIGGLASSINKAV-KISK-----GEYIFRMDADDISYSPRFDKQIRFMEENS--LDF 115

Db 64 AQAQNSGLISLNLGDELAKSGGGEYIARTDADIASPGWIEKIVGEMEDREIIM 123  
QY 116 SATLIELIDOK-GNLVYKQRESNKIYLT---NDIRKMLNRLSLAHPWCYKKVFDKL 170  
Db 124 GAWLEVLSEKEDGNRLARHKGKIMKPPRHEDIAAFPFGNPIHNNMTIMKRSVIDGG 183  
QY 171 MGY---RDLVPEVDYDF---AIRGALADFKIGLNLKVLQYRLNENGISQTNKFKQYIY 223  
Db 184 LRYDTERDM--AEDYQFWYDVSKIGRLAYP-----EALVYKRLHANQVSSKHSVRQH-- 234  
QY 224 SALLQDFYKESYIDITKITNYFOEYIKRYT---QOELSKYFEL--KSTPSITIRKL 277  
Db 235 -EIAQGIQK-----TARNDFLOSMGFKTRFDSLEYRQTKAAAYELPEKDLPEEDFERA 286  
QY 278 YICLYLYFK 286  
Db 287 RRFLYQCFK 295

## RESULT 7

US-08-683-426-11  
; Sequence 11, Application US/08683426  
; Patent No. 5705367  
; GENERAL INFORMATION:  
; APPLICANT: Gotschlich, Emil C.  
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF  
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/683,426  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/312,387  
; FILING DATE: September 26, 1994  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-095B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEEX: 133521  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 348 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-683-426-11

Query Match 13.2%; Score 210.5; DB 1; Length 348;

Best Local Similarity 27.8%; Pred. No. 2.3e-11;

Matches 86; Conservative 55; Mismatches 123; Indels 45; Gaps 15;

QY 4 SIMSVNPEPLNYVDSDVSIINQTLTDFEFLIVIDNSRQDLKQFLTEYSVVDNRKIL 63  
Db 6 SVLICAYNVE-KYFAGSILAAVAVNQTWRNDILITVDGSDTGLT-AIAKDFQKRSRIKIL 63

QY 64 LNEENIGLASSLNKAV-KISK-----GEYIFRMADADISYPSRFQKIRMEENS--LDF 115  
Db 64 AQAQNSGLISLNLGDELAKSGGGEYIARTDADIASPGWIEKIVGEMEDREIIM 123  
QY 116 SATLIELIDOK-GNLVYKQRESNKIYLT---NDIRKMLNRLSLAHPWCYKKVFDKL 170  
Db 124 GAWLEVLSEKEDGNRLARHKGKIMKPPRHEDIAAFPFGNPIHNNMTIMKRSVIDGG 183  
QY 171 MGY---RDLVPEVDYDF---AIRGALADFKIGLNLKVLQYRLNENGISQTNKFKQYIY 223  
Db 184 LRYDTERDM--AEDYQFWYDVSKIGRLAYP-----EALVYKRLHANQVSSKHSVRQH-- 234  
QY 224 SALLQDFYKESYIDITKITNYFOEYIKRYT---QOELSKYFEL--KSTPSITIRKL 277  
Db 235 -EIAQGIQK-----TARNDFLOSMGFKTRFDSLEYRQTKAAAYELPEKDLPEEDFERA 286  
QY 278 YICLYLYFK 286  
Db 287 RRFLYQCFK 295

## RESULT 8

US-08-683-458-3  
; Sequence 3, Application US/08683458  
; Patent No. 5798233  
; GENERAL INFORMATION:  
; APPLICANT: Gotschlich, Emil C.  
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF  
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/683,458  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/312,387  
; FILING DATE: September 26, 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-095A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEEX: 133521  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 348 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-683-458-3

Query Match 13.2%; Score 210.5; DB 1; Length 348;

Best Local Similarity 27.8%; Pred. No. 2.3e-11;

Matches 86; Conservative 55; Mismatches 123; Indels 45; Gaps 15;

QY 4 SIMSVNPEPLNYVDSDVSIINQTLTDFEFLIVIDNSRQDLKQFLTEYSVVDNRKIL 63  
Db 6 SVLICAYNVE-KYFAGSILAAVAVNQTWRNDILITVDGSDTGLT-AIAKDFQKRSRIKIL 63





Sequence 3, Application US/09353412  
Patent No. 6342382  
GENERAL INFORMATION:  
APPLICANT: Gotschlich, Emil C.  
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF  
OLIGOSACCHARIDES, AND GENES ENCODING THEM  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/333,412  
FILING DATE: 15-Jun-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/312,387  
FILING DATE: July 7, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-095  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 345-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 amino acids  
TYPE: amino acid





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2004, 03:30:59 ; Search time 101 Seconds  
(without alignments)

5159.395 Million cell updates/sec

Title: US-09-900-038a-2

Perfect score: 939  
Sequence: 1 atgaattatagatcatcatat.....gaggagagaacaagtgac 939

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/6C\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfltest1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	168.4	17.9	US-08-916-421B-1	Sequence 1, Appl
2	73.6	7.8	US-09-557-884-1	Sequence 1, Appl
3	73.6	7.8	US-09-643-980A-1	Sequence 1, Appl
4	69.8	7.4	US-09-453-702B-121	Sequence 12, App
5	64.8	6.9	US-09-495-406-14	Sequence 14, Appl
6	64.8	6.9	US-09-495-406-11	Sequence 1, Appl
7	63.8	6.8	US-09-495-406-11	Sequence 11, Appl
8	63.8	6.8	US-09-495-406-11	Sequence 1, Appl
9	63.4	6.8	US-09-134-000C-1681	Sequence 16, Appl
10	60.4	6.4	US-09-495-406-16	Sequence 16, Appl
11	58.8	6.3	US-09-134-000C-1682	Sequence 16, Appl
12	58.6	6.2	US-08-961-527-155	Sequence 15, App
13	56.6	6.0	US-08-487-826B-13	Sequence 13, Appl
14	55.6	5.9	US-08-487-826B-13	Sequence 13, Appl
15	51	5.4	US-08-998-416-534	Sequence 53, App
16	50.8	5.4	US-09-489-039A-1024	Sequence 1024, App
17	50.8	5.4	US-10-204-708-60	Sequence 60, Appl
18	50	5.3	US-08-998-416-287	Sequence 287, App
19	50	5.3	US-08-213-415B-3	Sequence 3, Appl
20	49.4	5.3	US-08-332-463-14	Sequence 14, Appl
21	49.4	5.3	US-08-961-527-71	Sequence 71, Appl
22	49.4	5.3	US-09-901-988-1	Sequence 1, Appl
23	48.8	5.2	US-08-257-073-12	Sequence 12, Appl
24	48.8	5.2	US-08-257-073-8	Sequence 8, Appl
25	48.6	5.2	US-08-998-416-186	Sequence 186, App
26	48.4	5.2	US-08-956-171B-182	Sequence 182, App
27	47.6	5.1	US-08-961-083-195	Sequence 195, App

28	47.6	5.1	844	4	US-09-536-784-195	Sequence 195, App
29	47.6	5.1	7304	4	US-10-204-708-44	Sequence 44, Appl
30	47.4	5.0	837	3	US-08-998-416-288	Sequence 288, App
31	47.4	5.0	6306	4	US-10-204-708-50	Sequence 50, Appl
32	46.6	5.0	640681	4	US-09-790-988-1	Sequence 1, Appl
33	46.4	4.9	2165	4	US-09-107-532A-3235	Sequence 3235, App
34	46.4	4.9	3701	3	US-08-845-258-10	Sequence 10, Appl
35	46	4.9	3701	3	US-08-990-571-10	Sequence 10, Appl
36	46	4.9	3701	4	US-08-723-142A-10	Sequence 10, Appl
37	46	4.9	3701	4	US-09-528-708A-10	Sequence 10, Appl
38	46	4.9	3701	4	US-09-569-098A-10	Sequence 10, Appl
39	45.6	4.9	636	3	US-08-998-416-1137	Sequence 1137, App
40	45.6	4.9	665	2	US-08-883-795A-36	Sequence 36, Appl
41	45.6	4.9	984	4	US-09-107-532A-2527	Sequence 2527, App
42	45.6	4.9	6152	3	US-08-973-462-1	Sequence 1, Appl
43	45.6	4.9	8920	3	US-08-446-855A-1	Sequence 1, Appl
44	45.6	4.9	8920	3	US-09-150-741-1	Sequence 1, Appl
45	45.6	4.9	15598	4	US-08-956-171B-82	Sequence 82, Appl

#### ALIGNMENTS

```

RESULT 1
US-08-916-421B-1/c
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
APPLICANT: Bult et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
Patent No. 6503729
FILE REFERENCE: PB275
CURRENT APPLICATION NUMBER: US/08/916,421B
CURRENT FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc_feature
LOCATION: (28222)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g

```

```

OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (10398)..(10398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312893)..(312893)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g

```

```

NAME/KEY: misc_feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779576)..(779576)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1139881)..(1139881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

```

```

Query Match 17.9%; Score 168.4; DB 4; Length 1664976;
Best Local Similarity 55.2%; Pred. No. 11e-27;
Matches 363; Conservative 5; Mismatches 281; Indels 9; Gaps 2;

QY 17 TTATGCGTATATATAGACCCCTTAAATTATGAGACATTCAGTAGATCTATATTAA 76
Db 475813 TAATGCAACATACACGAAACGAAATAATTAAGAACTATGATCAATTTVA 475754
QY 77 ATCAAGCCTTACTGATTTTGGCTTATATTTGTCTATTATATCCAAAGTAGAGTGAT 136
Db 475753 ATCAACATTTAAGATTGTAATTTATATCGYTTATATATCCAAATATATAAAG 475694
QY 137 TAAAGCAATCTTAAGAAATATTCAGTTGTAGATATGAATTAATAATCTGCTTAAG 196
Db 475693 CAGAGCAATTTAATTAAGATATCAACAGAAATATAATTTTATATAAATG 475634
QY 197 AAGAAATATGCTTATGATCAAGTTTGAACAACAGCGTGAATTTCTAAGGAGAA 256

```

```

DB 475613 AAGAAATTTGATGATGAGGAGCCGATAGAAATTAAGCTGTATATATAGCAAGGGAGT 475574
QY 257 ATATTTTGAATGATGCTGATGATATTTTATATCCAGTAGATTTGATAGCAATTC 316
DB 475573 ATATTTGCCATATTTAGACCTGATGATATTTGATATCCATTAAGATTAAGAAACAAATTTA 475514
QY 317 GTTTATAGAGAGAAATTTCA---TTGATTTCTCGCAACCTATATGATTTGATAGACC 373
DB 475513 AATATATGAGAAATTAATAGACATGATTTATTTAGTGGGTTTATTTATTTGATG 475454
QY 374 AAAAAAGAAATTTAGTATATTAACACAGAGAAAGTAAATTAATTTACTTAACTATATGATA 433
DB 475453 AAAAAAGAAATTTA-----TTTAAAAAGATTTAAGCCAGAAAAATTAATTAAGAAA 475400
QY 434 TACGGAATATGTTATGTAATGATCTATATCTGCCCACCAAGCGTGCTTAAGAAAAGA 493
DB 475399 TTAATAAATATTTTTCATAAGACATTTACTTTTCAATCCATCTATGATGTTAAAGATA 475340
QY 494 AAGTTTTCATTAAGTTAATGGATATAGAGATTTTAGTACTGTTGAAGATTATGATTTTG 553
DB 475339 AATCTTAAGAAATTTAATAATGATGAGAAATTAATTAGGCTTCAAGACTACATTTT 475280
QY 554 CATATAGAGAGCTCTGCTGATTTTCAAAATCGGCTTCAATTAAGTACTTTTACGCT 613
DB 475279 GGATTAGATGATAGCCCAATGATTAACAAATTTGACATTAATGAGAAATTTTATTAAT 475220
QY 614 ATAGATTAACGAGATGGAATATACCAAAACCAATTAAGTTAAGCAATATATTTACTG 671
DB 475219 ATAGATTTCCAAATTAAGACATTTTATTAAGCAGAAATCAAAAAATTAATATTC 475162

```

```

RESULT 2
US-09-557-884-1
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
;
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB186P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

```

```

; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1
Query Match 7.8%; Score 73.6; DB 4; Length 1830121;
Best Local Similarity 54.9%; Pred. No. 5.6e-07;
Matches 167; Conservative 0; Mismatches 134; Indels 3; Gaps 1;
QY 30 TAATGACCTTTAAATTATGATGAGATTCAGTAGATCTATATTAATCAACGCTTAC 89
DB 1646128 TTTAAGCTGAGCATATATATGATGAAGACATTCATCATTAATTAATCAAGCTATGA 1646187
QY 90 TGAATTTGATCATATATTTGCTATGATTAATCCAGTAGAGGATTTAAACGATTCCT 149
DB 1646188 AATCTTGAATAATATATGATATCAATGATGTTCAACAGATTGACTTGTCTCA--TTT 1646244
QY 150 AACAGAAATTTAGTTGATGATATATGAATTAATAATCTTGTAAATGAAGAAATATTGG 209
DB 1646245 AGAAGAAATATCTAAATTAATGATTAAGATTAATAATTAATCAAGTAATTAATTTAG 1646304
QY 210 TTTAGATCAAGTTTGAACAAAGCGTGAAATTTCTAAGGAGAAATATATTTTGAAAT 269
DB 1646305 GTTCATTAATTTCTTGAATATAGCCCTTGCTGTTTTCAGGTAATATTTTGCAAGAA 1646364
QY 270 GGATGCTGATGATATTTTCATATCCAGTAGATTTGATAAGCAAAATCGTTTATGAGAGA 329
DB 1646365 GGATGCTGATGATATAGCTTAACCAATCGTGATGAGAAATATGTTACTATCTGAGAA 1646424
QY 330 AAT 333
DB 1646425 AAT 1646428

```

```

RESULT 3
US-09-643-990A-1
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Owen White
; Hamilton O. Smith
; J. Craig Venter
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
;
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB186P1C1
; TELECOMMUNICATION INFORMATION:

```

TELEPHONE: 301-610-5790  
 TELEFAX: 310-309-8439  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1830121 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-643-990A-1

Query Match 7.8%; Score 73.6; DB 4; Length 1830121;  
 Best Local Similarity 54.9%; Pred. No. 5.6e-07;  
 Matches 167; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

QY 30 TAAAGACCTTAAATATGAGAGATTCAGTGAATCTATTAATCAAGCGCTTAC 89  
 DB 1646128 TTAATACGCTGAGCAATATATGATGAAGCAATTCATCAATTAATCAAGCTTATGA 1646187  
 QY 90 TGATTTGAGTTCATTAATTTGCTATTGATTAATCCAGTAGAGGTGATTTAAGCAATTCCT 149  
 DB 1646188 AAATCTGAATTTATGATTAATCAATGATGTTCAAGATTTGACTTGTCTCA---TTT 1646244  
 QY 150 AACGAATATTCAGTGTAGTATTAATGAATTAATTTCTGCTTAATGAAGAAATTTGG 209  
 DB 1646245 AGAAGAAATATCTTAATTAATGAATTAATTAATTCATGATTAATTAATTTAGG 1646304  
 QY 210 TTAAGCATCAAGTTTGAACAAGCGGTGAATTTCTAAGGAGAAATATATTTTGAAT 269  
 DB 1646305 GTTCATAATCTTTGAAATATAGCGCTTGCTGTTTTCAGGTAAATTTTGCAAGAA 1646364  
 QY 270 GGAATGCTGATATATTTTCAATTCAGATGAATTTGATTAAGCAATTTGTTTATGAGGA 329  
 DB 1646365 GGAATGCTGATATATTTGATTAATCAATTCATGATTAAGCAATTAATTTAGTACCTATCTGGAGAA 1646424  
 QY 330 AAAT 333  
 DB 1646425 AAAT 1646428

RESULT 4  
 US-09-453-702B-121/C  
 Sequence 121, Application US/09453702B  
 Patent No. 6365723  
 GENERAL INFORMATION:  
 APPLICANT: Blatner, Frederick R.  
 Inventor: Blatner, Frederick R.  
 Perna, Nicole T.  
 Plunkett, Guy  
 Welch, Rod  
 TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157  
 NUMBER OF SEQUENCES: 265  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: Quarles & Brady  
 STREET: 1 South Plinkney Street  
 CITY: Madison  
 STATE: WI  
 COUNTRY: US  
 ZIP: 53701-2113  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Word Perfect 8.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/453,702B  
 FILING DATE: 03-Dec-1999  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/110,955  
 FILING DATE: 04-Dec-1998  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Seay, Nicholas J.

REGISTRATION NUMBER: 27386  
 REFERENCE/DOCKET NUMBER: 960296.95017  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (608) 251-5000  
 TELEFAX: (608) 251-9166  
 INFORMATION FOR SEQ ID NO: 121:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 14187  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 SEQUENCE DESCRIPTION: SEQ ID NO: 121:  
 US-09-453-702B-121

Query Match 7.4%; Score 69.8; DB 4; Length 14187;  
 Best Local Similarity 52.7%; Pred. No. 1.6e-06;  
 Matches 175; Conservative 0; Mismatches 154; Indels 3; Gaps 1;

QY 58 TCAGTAGATCTATTAATTAATCAAGCTTACTGATTTTGAGTTCATTAATGTCATTTGAT 117  
 DB 13859 TCAGTAGATCTATTAATTAATCAATCTTAATCAAGATTTTG---TTTGTATATCATTTGAC 13803  
 QY 118 AATCCAGTAGAGGTGATTTAAGCAATTTCTTAACAGATATTCAGTTGTAGATTAATGA 177  
 DB 13802 GATTTGACACCGATGATACATTTTTCATTAATCAACAGTCATCAAAAACATCAGAAA 13743  
 QY 178 ATTAATAATCTGCTTAATGAAGAAATATTTGCTTTAGCATCAAGTTTGAACAAAGCGGTG 237  
 DB 13742 ATTAATAATCTGCTTAATGAAGCAATTTAGGTGTTGAGAAAGTCGAAATTAATGAATA 13683  
 QY 238 AAAATTTCTAAGGAGGAATATATTTTGAATGATGCTGATGATTTTCATTCAGTCAAGT 297  
 DB 13682 GAAATGCGCCAGCGGAATATATTTCTTTTGTATGCGATGATTTTGCGACAGAGAA 13623  
 QY 298 AGATTTGATTAAGCAATTTGTTTATGAGAGAAATTCATGATTTGTCGCAATCTCTA 357  
 DB 13622 AATTAAGAGCGTCAATGCAAGTGTAAATATGATGATGATGATGATGATGATGATGAT 13563  
 QY 358 ATGAATGATGATGACCAAAAGCAATTTAGT 389  
 DB 13562 TATTAATGATGATGATGACCAAAAGCAATTTAGT 13531

RESULT 5  
 US-09-495-406-14  
 Sequence 14, Application US/09495406  
 Patent No. 6503744  
 GENERAL INFORMATION:  
 APPLICANT: Gilbert, Michel  
 APPLICANT: Wakarchuk, Warren W.  
 APPLICANT: National Research Council of Canada  
 TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 FILE REFERENCE: 019633-000110US  
 CURRENT APPLICATION NUMBER: US/09/495,406  
 CURRENT FILING DATE: 2000-01-31  
 PRIOR APPLICATION NUMBER: US 60/118,213  
 PRIOR FILING DATE: 1999-02-01  
 NUMBER OF SEQ ID NOS: 35  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 14  
 LENGTH: 906  
 TYPE: DNA  
 ORGANISM: Campylobacter jejuni  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (1)..(906)  
 OTHER INFORMATION: beta-1,3-galactosyltransferase from C. jejuni strain  
 OTHER INFORMATION: OHA384 (ORF 6a of LOS biosynthesis locus)  
 US-09-495-406-14

Query Match 6.9%; Score 64.8; DB 4; Length 906;

Best Local Similarity 51.8%; Pred. No. 1.2e-05;  
Matches 172; Conservative 0; Mismatches 157; Indels 3; Gaps 1;

QY 1 ATGAATTAATGATCATTATGTCGGTATATATAGAGCTTTAAATTAATGAGAGATTCA 60  
DB 1 ATGTTAAATTTTCATCATCTTACCACTTAATATGTAACAATATATATGCAAGGCA 60  
QY 61 GTAGAACTCTATATTAATCAACGCTTACGATTTGAGTCATTAATGTCATTGATAT 120  
DB 61 GTAGAACTCTATATTAATCAACGCTTACGATTTGAGTCATTAATGTCATTGATAT 120  
QY 61 ATAGAACTCTATATCAATCAAGCTTTTAAAGATATAG-...AATATTTGATGATGAT 117  
DB 121 CCAAGTAGAGGTGATTTAAAGCAATCTTAAACAGATATTCAGTTGATATAGATA 180  
QY 118 TGTGAAATGATATATATATATATATATAGCCAAAGATATCTTAAAAAACAAAGATA 177  
DB 181 AAAATCTTGCTTAATAGAGAAATATTTGTTAGATCAAGTTTGAACAAAGCGGTGAA 240  
QY 178 AAAATTAATCCAAATGAAAAAACTTAGGCTTTTAAAGCAAGATATGAAAGGTGAA 237  
DB 241 ATTTTAAGGAGAAATATATTTTGAATGATGCTGATATATTCATATCAAGTGA 300  
QY 238 GTAGCAAACTCTCCCTTAATATATGTTTAAATCTGATGATTTTGAAGTAAATGCT 297  
DB 301 TTTGATTAAGCAATTCGTTTATGAGGAAAA 332  
QY 298 TGTGAAGGTGTATATAAAATTTTATGATGAAAC 329

RESULT 6  
US-09-495-406-1/c  
Sequence 1, Application US/09495406  
Patent No. 6503744  
GENERAL INFORMATION:  
APPLICANT: Gilbert, Michel  
APPLICANT: Wakarchuk, Warren W.  
APPLICANT: National Research Council of Canada  
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
FILE REFERENCE: 019633-000110US  
CURRENT FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: US 60/118,213  
PRIOR FILING DATE: 1999-02-01  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 11474  
TYPE: DNA  
ORGANISM: Campylobacter jejuni  
FEATURE:  
OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384  
US-09-495-406-1

Query Match 6.9%; Score 64.8; DB 4; Length 11474;  
Best Local Similarity 51.8%; Pred. No. 1.9e-05;  
Matches 172; Conservative 0; Mismatches 157; Indels 3; Gaps 1;

QY 1 ATGAATTAATGATCATTATGTCGGTATATATAGAGCTTTAAATTAATGAGAGATTCA 60  
DB 5959 ATGTTAAATTTTCATCATCTTACCACTTAATATGTAACAATATATATGCAAGGCA 5900  
QY 61 GTAGAACTCTATATTAATCAACGCTTACGATTTGAGTCATTAATGTCATTGATAT 120  
DB 5899 ATAGAAAGCTGATCAATCAAGCTTTTAAAGATATAG--AATATTTGATGATGAT 5843  
QY 121 CCAAGTAGAGGTGATTTAAAGCAATCTTAAACAGATATTCAGTTGATATAGATA 180  
DB 5842 TGTGAAATGATATATATATATATATATAGCCAAAGATATCTTAAAAAACAAAGATA 5783  
QY 181 AAAATCTTGCTTAATAGAGAAATATTTGTTAGATCAAGTTTGAACAAAGCGGTGAA 240  
DB 5782 AAAATTAATCCAAATGAAAAAACTTAGGCTTTTAAAGCAAGATATGAAAGGTGAA 5723

QY 241 ATTTTAAGGAGAAATATATTTTGAATGATGCTGATGATATTTCAATATCAAGTGA 300  
DB 5722 GTAGCAAACTCTCCCTTAATATATGTTTAAATCTGATGATTTTGAAGTAAATGCT 5663  
QY 301 TTTGATTAAGCAATTCGTTTATGAGGAAAA 332  
DB 5662 TGTGAAGGTGTATATAAAATTTTATGATGAAAC 5631

RESULT 7  
US-09-495-406-1  
Sequence 1, Application US/09495406  
Patent No. 6503744  
GENERAL INFORMATION:  
APPLICANT: Gilbert, Michel  
APPLICANT: Wakarchuk, Warren W.  
APPLICANT: National Research Council of Canada  
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
FILE REFERENCE: 019633-000110US  
CURRENT FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: US 60/118,213  
PRIOR FILING DATE: 1999-02-01  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 11  
LENGTH: 1170  
TYPE: DNA  
ORGANISM: Campylobacter jejuni  
FEATURE:  
OTHER INFORMATION: glycosyltransferase from C. jejuni strain OH4384  
US-09-495-406-11

Query Match 6.8%; Score 63.8; DB 4; Length 1170;  
Best Local Similarity 56.8%; Pred. No. 2.1e-05;  
Matches 138; Conservative 0; Mismatches 102; Indels 3; Gaps 1;

QY 42 AAATATGATGAGATTCAGTATGATATATTAATTAATCAACGCTTACGATTTGAGTT 101  
DB 42 AAATATTAATGAGAAATGTTATGATAGCGTTATCAATCAAACTTATATCACTTAAGAAAT 101  
QY 102 CATATGTCATGATATATCAAGTATGAGGTATTTAAAGCAATCTTAAACAGATATTC 161  
DB 102 CATCTTGTCAATGATGATGACAGATGACACTCATCAATATGCAAAAATATATAC 161  
QY 162 AGTTGATATATATGATATTAATTAATCTTCTTAATGAGAAATATTTGTTAGCATCAAG 221  
DB 162 CTTAAAGATTAAGAAATATCTCTTT--TTGATTAAGAAAAATGGGGGTTTAAGTTACAGC 218  
QY 222 TTGAAACAACGCGTGAATTTCTAAGGAGATATATTTTATGATGATGATGATGATGA 281  
DB 219 TGAATAATATGATATGATATATCTTTAGCGGGAATATTAATTAATAAACAACCTCAAC 278  
QY 282 TAT 284  
DB 279 TAT 281

RESULT 8  
US-09-495-406-1  
Sequence 1, Application US/09495406  
Patent No. 6503744  
GENERAL INFORMATION:  
APPLICANT: Gilbert, Michel  
APPLICANT: Wakarchuk, Warren W.  
APPLICANT: National Research Council of Canada  
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
FILE REFERENCE: 019633-000110US  
CURRENT APPLICATION NUMBER: US/09/495,406

CURRENT FILING DATE: 2000-01-31  
 PRIOR APPLICATION NUMBER: US 60/118,213  
 PRIOR FILING DATE: 1999-02-01  
 NUMBER OF SEQ ID NOS: 35  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 1  
 LENGTH: 11474  
 TYPE: DNA  
 ORGANISM: Campylobacter jejuni  
 FEATURE:  
 OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384  
 OTHER INFORMATION: including LOS biosynthesis locus  
 US-09-495-406-1

Query Match 6.8%; Score 63.8; DB 4; Length 11474;  
 Best Local Similarity 56.8%; Pred. No. 3.1e-05;  
 Matches 138; Conservative 0; Mismatches 102; Indels 3; Gaps 1;

QY 42 AATTATGATGAGATTCAGTATGATCTATATTAACCAACGCTTACGATTTGAGTT 101  
 DB AAAATATTAGAGAAATGTTAGTACCGTTATCAATCAACTTAACTAATTAAGAAAT 2886  
 QY 102 CATAATGTCATGATATCAAGTATGAGTATTAAGCAATTTCTTAACAGAAATTTTC 161  
 DB 2887 CATCTTGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2946  
 QY 162 AGTTGTAGATTAATGAAATTAATTAATTTCTGCTTAATGAAAGAAATTTGCTTACATCAAG 221  
 DB 2947 CTAAAGATTAATGAAATTAATTAATTTCTGCTTAATGAAAGAAATTTGCTTACATCAAG 3003  
 QY 222 TTGAACAAGCGGTGAAATTTCTAAGGAGAAATTAATTTTGAATGATGCTGATCA 281  
 DB 3004 TAGAAATTAATGATTAATTAATTTCTAAGGAGAAATTAATTTTGAATGATGCTGATCA 3063  
 QY 282 TAT 284  
 DB 3064 TAT 3066

RESULT 9  
 US-09-134-000C-1681  
 Sequence 1681, Application US/09134000C  
 Patent No. 6617156  
 GENERAL INFORMATION:  
 APPLICANT: Lynn Doucette-Stamm et al  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 032796-032  
 CURRENT APPLICATION NUMBER: US/09/134,000C  
 CURRENT FILING DATE: 1998-08-13  
 PRIOR APPLICATION NUMBER: US 60/055,778  
 PRIOR FILING DATE: 1997-08-15  
 NUMBER OF SEQ ID NOS: 6812  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 1681  
 LENGTH: 3171  
 TYPE: DNA  
 ORGANISM: Enterococcus faecalis  
 US-09-134-000C-1681

Query Match 6.8%; Score 63.4; DB 4; Length 3171;  
 Best Local Similarity 50.8%; Pred. No. 3.1e-05;  
 Matches 151; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 12 TATCATTAATGCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 71  
 DB 1698 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1757  
 QY 72 AATTAATCAAGCTTACTGATTTGATGATCAATTTGATGATCAATTTGATGATCAAGTAGAG 131  
 DB 1758 ACAAATCAAGCTTACTGATTTGATGATCAATTTGATGATCAATTTGATGATCAAGTAGAG 1817  
 QY 132 TGAATTAAGCAATTTCTTAACAGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 191

DB 1818 ACATATTAGCAATGCTGATGAAATTAATTAAGAACTGATCAACGAATCAAGTTATTA 1877  
 QY 192 TATGAGAAATTAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 251  
 DB 1878 TCGGAGAAATTAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1937  
 QY 252 AGAATTAATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 308  
 DB 1938 CGATTTTATGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1994

RESULT 10  
 US-09-495-406-16  
 Sequence 16, Application US/09495406  
 Patent No. 6503744  
 GENERAL INFORMATION:  
 APPLICANT: Gilbert, Michel  
 APPLICANT: Wakarchuk, Warren W.  
 APPLICANT: National Research Council of Canada  
 TITLE OF INVENTION: Campylobacter glycosyltransferases for Biosynthesis of  
 TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 FILE REFERENCE: 019633-000110US  
 CURRENT APPLICATION NUMBER: US/09/495,406  
 CURRENT FILING DATE: 2000-01-31  
 PRIOR APPLICATION NUMBER: US 60/118,213  
 PRIOR FILING DATE: 1999-02-01  
 NUMBER OF SEQ ID NOS: 35  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 16

QY 912  
 DB 912  
 TYPE: DNA  
 ORGANISM: Campylobacter jejuni  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (1)..(912)  
 OTHER INFORMATION: Campylobacter glycosyltransferase B (CgtB) beta-1,3  
 OTHER INFORMATION: galactosyltransferase from C. jejuni serotype O:2  
 OTHER INFORMATION: (strain NCTC 11168)  
 US-09-495-406-16

Query Match 6.4%; Score 60.4; DB 4; Length 912;  
 Best Local Similarity 51.2%; Pred. No. 0.00011;  
 Matches 167; Conservative 0; Mismatches 156; Indels 3; Gaps 1;

QY 1 ATGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60  
 DB 1 ATGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60  
 QY 61 GTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 120  
 DB 61 GTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 117  
 QY 121 CCAGTAGAGTGAATTAAGCAATTTCTTAACGAAATTTCAAGTATTTCAAGTATTAATTA 180  
 DB 118 TGTGTAATGATTAAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 177  
 QY 181 AAAATTTGCTTAATGAAGAAATTAATTTGCTTAATGAAGTATTAATGAAGTATTAATGA 240  
 DB 178 AAAATTTGCTTAATGAAGAAATTAATTTGCTTAATGAAGTATTAATGAAGTATTAATGA 237  
 QY 241 ATTTTGAAGGAAATTAATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 DB 238 GTGCAATTTCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 297  
 QY 301 TTGATTAAGCAATTTGCTTTATGA 326  
 DB 298 TCGAAGAAATGATTAATTAATTTTGA 323

RESULT 11  
 US-09-134-000C-1682  
 Sequence 1682, Application US/09134000C

Patent No. 6617156  
GENERAL INFORMATION:  
APPLICANT: LYNN Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
FILE REFERENCE: 032796-032  
CURRENT APPLICATION NUMBER: US/09/134,000C  
PRIOR FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/055,778  
NUMBER OF SEQ ID NOS: 6812  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1682  
LENGTH: 1602  
TYPE: DNA  
ORGANISM: Enterococcus faecalis  
US-09-134-000C-1682

Query Match 6.3%; Score 58.8; DB 4; Length 1602;  
Best Local Similarity 49.1%; Pred. No. 0.00027;  
Matches 156; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

QY 19 ATGTCGATATATATAGGCTTTAATATATGAGATTCAGTAATCTATATATAT 78  
DB 595 ATGCCAGTATATATGTTGAGAAAATGCGTTCGTTGTGATTTATTTGAT 654  
QY 79 CAACGCTTACTGATTGATTCATATATGCTATATATCAAGTAGAGGATTTA 138  
DB 655 CAAGCTATACCTACCTGGAATATATATGAGCCGATGAGCCCTCAAGATCTATGTC 714  
QY 139 AAGCAATCTTAAAGATATATTCAGTGTAGATATATATATATATATCTTCTTATGAA 198  
DB 715 AAAAATTTTAAAGAGATCAAGCATGATGAGGATTCGCTCTTCTTGTGAA 774  
QY 199 GAAATATTTGTTTACATCAAGTTGAACAAAGCGGTGAAAATTTCTTAAAGAGATAT 258  
DB 775 CAAAATGTCATATTTCCAGACCAACCACTGCTTTGCGATGCTACCGAGAAATTT 834  
QY 259 ATTTTGAATGATCTGATATATTTTATATCAAGATATGATTAAGCAATTCGT 318  
DB 835 GTCCGCTTACTAGACACATATGATATAGCATCAATGCTTTTATGAGTGTAA 894  
QY 319 TTTATGAGCAAAATTTCA 336  
DB 895 GTGCTGAATGAAAACCCA 912

## RESULT 12

US-08-961-527-155/c  
Sequence 155, Application US/08961527  
Patent No. 6420135

## GENERAL INFORMATION:

APPLICANT: Charles Kunsch  
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 391

CORRESPONDENCE ADDRESS:  
ADDRESS: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue

CITY: Rockville  
STATE: Maryland  
COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,527

FILING DATE:  
CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8512  
TELEFAX: (301) 309-8504  
INFORMATION FOR SEQ ID NO: 155:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6474 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-961-527-155

Query Match 6.2%; Score 58.6; DB 4; Length 6474;  
Best Local Similarity 45.2%; Pred. No. 0.00039;  
Matches 303; Conservative 0; Mismatches 359; Indels 9; Gaps 2;

QY 44 ATATGTGAGAGATTCAGTGAATCTATATTAATCAACGCTTACTGATTTGAGTTCA 103  
DB 5378 AATCTCGATTAGAGCTGTGAAAAGTTGTTTACATCAAACTTATTAATAATGAAAGTTA 5319  
QY 104 TAATGTCAATTGATATATCAAGTAGAGGATTTTAAAGCAATCTTAAAGATAT 160  
DB 5318 TAATATGATGAT 5259  
QY 161 CAGTTGAT 220  
DB 5258 GCTATACCAT 5199  
QY 221 GTTTGACAAAGGAGGTAATTTCTTAAAGGAGATATATATATATATATATATAT 280  
DB 5198 CCAGAACAT 5139  
QY 281 ATATTTGAT 340  
DB 5138 ATGAAAT 5079  
QY 341 ATTTCTGAGCACTTAT 5078  
DB 5078 AGAATTTAGC-----TTTATGTTGTTATGATATATATATATATATATATAT 5025  
QY 401 GAGAAAT 460  
DB 5024 GAGAAAGAGAGAGAGAGAGATTTTGTGAATCCCTTTGTTCAATGATCAATA 4965  
QY 461 TACTTGCCCAACCAAGTGTGCTTAAAGAAAGTTTCCGATTAATGAGATATA 520  
DB 4964 TAGCAGGTAGCTCAATTTGTGTGTTAAAGAGGCTGATGATTAATTAATGTTTG 4905  
QY 521 GAGATTAGTACCTGTTGAAGATATATATATATATATATATATATATATATAT 580  
DB 4904 AGAAAT 4845  
QY 581 AAATGCGCTTATCAATTAAGTACTTTTACATATATATATATATATATATAT 640  
DB 4844 AAATGAT 4785  
QY 641 AAACCAAT 700  
DB 4784 TCACTGAGTACCAAAACAAATTAATGACATGAAATATATATATATATATAT 4725  
QY 701 AATCTTATAT 711  
DB 4724 AATATTTTAT 4714

## RESULT 13

US-08-487-826B-13/c  
Sequence 13, Application US/08487826B  
Patent No. 5993827

GENERAL INFORMATION:  
 APPLICANT: Sim, Kim L.  
 APPLICANT: Chitnis, Chetan  
 APPLICANT: Miller, Louis H.  
 APPLICANT: Peterson, David S.  
 APPLICANT: Su, Xin-zhaun  
 APPLICANT: Wellens, Thomas E.  
 TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
 TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
 NUMBER OF SEQUENCES: 45  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: Knodde Martens Olson & Bear  
 STREET: 620 Newport Center Drive 16th Floor  
 CITY: Newport Beach  
 STATE: California  
 COUNTRY: US  
 ZIP: 92660  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/487,826B  
 FILING DATE: 10-SEP-1993  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Israelien, Ned  
 REGISTRATION NUMBER: 29,655  
 REFERENCE/DOCKET NUMBER: NIH121.001CP1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 235-8550  
 TELEFAX: (619) 235-0176  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 19124 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: CDNA  
 HYPOTHEICAL: NO  
 ANTI-SENSE: NO  
 US-08-487-826B-13

Query Match 6.0%; Score 56.6; DB 2; Length 19124;  
 Best Local Similarity 43.6%; Pred. No. 0.0013;  
 Matches 363; Conservative 0; Mismatches 459; Indels 11; Gaps 2;

QY 11 GTATCATATGTCGGTATATATAGACCTTAAATATGAGAGATTCAAGTACATCTA 70  
 DB 18461 GTTAAAT 18402

QY 71 TATTAATCAACAGCTTACTGATTTTGATCATATGTCATATGATTAATCCAGTACAG 130  
 DB 18401 AAGTGAAGCTATGTTTATATATTTTATATATATATATATATATATATATATAT 18342

QY 131 GTGATTTAAGCAATCTTAAACAGATATTCAGTTGATGATATATATATATATATAT 190  
 DB 18341 TTTTGTGTTGATATAGATATATATATATATATATATATATATATATATATAT 18282

QY 131 TTAATGAAGAAATTTGTTTGAAGTCAAGTTTGACAAACCGGTGAAATTTCTAAG 250  
 DB 18281 GAAATTAAGAAATTAAGTAAATTAAGAAATTAAGAAATTTGTTATATATATAT 18222

QY 251 GAGATATATATTTTGAATGATGCTGATATATTTTCATATCCAGTACATATAT 306  
 DB 18221 CAAGAAAT 18162

QY 307 AAGCAATTCGTTTATAGAGAAATTCATGATTTCTAGCAACTATATGATG 366  
 DB 18161 TTATTAATATATTTTATATAGATATATATATATATATATATATATATATAT 18102

QY 367 ATAGACCAAAAGAAATTTAGTATATATATATATATATATATATATATATATAT 426

DB 18101 GAGTAACAAAATTAATAACACATATATATATATATATATATATATATATAT 18042

QY 427 AATGATATACGAAGATGTTATGTAATGATCTATATCTTGGCCACCCAGCTGGCGTA 486

DB 18041 ATTATTAATATACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 17989

QY 487 AAAAAAGAAATTTGATATAGTAAATGAGATATAGAGATTTAGTACCTGTTGAAGATAT 546

DB 17988 AATCGAAATATAGAAAGATATATATATATATATATATATATATATATATATAT 17929

QY 547 GATTTTGCATATAGAGAGCTGTGCTGATTTTCAAAATCGGCTTACTCAATTAAGTACT 606

DB 17928 TATATTACGAAAAATGACAGAAAAACAAATTAACATATATATATATATATAT 17869

QY 607 TTACAGTATAGATTAAGAGAAATGAAATATATCAACAACCAATTAAGTTAAGCAATATAT 666

DB 17868 ATCAATATATAGAT 17809

QY 667 TACTGCTATTTTCAAGATTTTATATATATATATATATATATATATATATATATAT 726

DB 17808 ATGAAAAAATTAACCAATTAAGAAAAAAGAAAAAATCATGTTGCAATATATATAT 17749

QY 727 ACTAATTAATTTCAAGATATGATTAAGAAAAACCTATATCTACAGACAGAGCTCTAA 786

DB 17748 ATATATATATGATTTATTAACAAAAACATATTTTTCACAAATATACCAAAATTAAT 17689

QY 787 TATTTGACCTAAATCTACCCCTGATTTACTATTAAGAAACATATATATTTG 839

DB 17688 AATTAAT 17636

RESULT 14  
 US-08-487-826B-13  
 Sequence 13, Application US/08487826B  
 Patent No. 5993827  
 GENERAL INFORMATION:  
 APPLICANT: Sim, Kim L.  
 APPLICANT: Chitnis, Chetan  
 APPLICANT: Miller, Louis H.  
 APPLICANT: Peterson, David S.  
 APPLICANT: Su, Xin-zhaun  
 APPLICANT: Wellens, Thomas E.  
 TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
 TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
 NUMBER OF SEQUENCES: 45  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: Knodde Martens Olson & Bear  
 STREET: 620 Newport Center Drive 16th Floor  
 CITY: Newport Beach  
 STATE: California  
 COUNTRY: US  
 ZIP: 92660  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/487,826B  
 FILING DATE: 10-SEP-1993  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Israelien, Ned  
 REGISTRATION NUMBER: 29,655  
 REFERENCE/DOCKET NUMBER: NIH121.001CP1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 235-8550  
 TELEFAX: (619) 235-0176  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 19124 base pairs  
 TYPE: nucleic acid



STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: CDNA  
 HYPOHETICAL: NO  
 ANTI-SENSE: NO  
 US-08-487-826B-13

Query Match 5.9%; Score 55.6; DB 2; Length 19124;  
 Best Local Similarity 42.5%; Pred. No. 0.0021;  
 Matches 351; Conservative 0; Mismatches 474; Indels 1; Gaps 1;

QY 94 TTGAGTTCATTAATGTCATGATTAATCCAGTAGAGGTGATTTAAAGCAATCTTAAACA 153  
 DB 15377 TATGCTTTTAT 15436  
 QY 154 GAATATTCAGTTGATGATTAATGATTAATATCTGCTTATGAGAAATATGCTTTA 213  
 DB 15437 CAATTTTATTAATGAAAAAGAAAAATGAAATATATATATATATATATATATATAT 15496  
 QY 214 GCATCAAGTTGAAACAAAGCGGTGAAAATTTCTAAGGAGATATATTTTATGATGAT 273  
 DB 15497 AAAAAAAGAAAAAGAAAAAGAAAAATTTTAAAAAATATATATATATATATATAT 15556  
 QY 274 GCTGATGATTTCTATCCAGTAGATTTGATTAAGCAAAATCGTTTATGAGAGAAAT 333  
 DB 15557 ATAAATTTTATGATTAATGAAAAAGAAAAAGATTTTCAAAAAAATTTAAAAAATTT 15616  
 QY 334 TCATTTGATTTCTCAGCACTCTATAGATTTGATTAAGCAAAAAAGAAATTTATATAT 393  
 DB 15617 TAT 15676  
 QY 394 AAACACAGGAAAGATTAAT 453  
 DB 15677 AAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 15736  
 QY 454 AGATCTATATCTGCCCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 513  
 DB 15737 AAAT 15796  
 QY 514 GGATTTAGAGATTTGATTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 573  
 DB 15797 AAAT 15856  
 QY 574 GATTTCAAAATGCGCTTCTCATTAAGATCTTTAGATTTAGAGAGATTTAGAGATTT 633  
 DB 15857 AAAAAAAT 15916  
 QY 634 ATATCAAAACCAATTAAGATTTAGCAATATATATATATATATATATATATATATAT 693  
 DB 15917 AT 15976  
 QY 694 AAAGAAATATCTTATTTGATTAATCAAAATATCTATATATATATATATATATATAT 752  
 DB 15977 ACAT 16036  
 QY 753 AAAGAAAGCTTATCTCAGCAAGCTCTTAATATTTTGAAGCTAAATCTTACCCCTAG 812  
 DB 16037 TAT 16096  
 QY 813 TATTTCTATTAAGAAAT 872  
 DB 16097 AT 16156  
 QY 873 GAGTTAT 918  
 DB 16157 TACCTAAT 16202

RESULT 15  
 US-08-998-416-534/c  
 ; Sequence 534, Application US/08998416  
 ; Patent No. 6239264  
 ; GENERAL INFORMATION:

APPLICANT: Philippsen, Peter  
 APPLICANT: Pohlmann, Rainer  
 APPLICANT: Steiner, Sabine  
 APPLICANT: Mohr, Christine  
 APPLICANT: Wendland, Jürgen  
 APPLICANT: Knechtle, Philipp  
 APPLICANT: Redischung, Corinne  
 TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSWYPII  
 TITLE OF INVENTION: AND USES THEREOF  
 NUMBER OF SEQUENCES: 1152  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: No. 6239264aris Corporation  
 STREET: 3054 Cornwallis Road  
 CITY: Research Triangle Park  
 STATE: No. 6239264th Carolina  
 COUNTRY: USA  
 ZIP: 27709  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/998,416  
 FILING DATE: 24-DEC-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: CH 0016/97  
 FILING DATE: 31-DEC-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meigs, J. Timothy  
 REGISTRATION NUMBER: 38,241  
 REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 919-541-8587  
 TELEFAX: 919-541-8689  
 INFORMATION FOR SEQ ID NO: 534:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 854 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 ORIGINAL SOURCE:  
 ORGANISM: PAG1372UP  
 US-08-998-416-534

Query Match 5.4%; Score 51; DB 3; Length 854;  
 Best Local Similarity 45.1%; Pred. No. 0.012;  
 Matches 265; Conservative 0; Mismatches 320; Indels 2; Gaps 2;

QY 41 TAAATATGAGAGATTCGATGATTAATTAATCAAGCTTACTGATTTGAGCT 100  
 DB 649 TAAATATATAGAAACCAATAGATTAATGATTCATTAAGAAATATATTTGAGCA 590  
 QY 101 TCATTAATGTCATGATTAATCCAGTAGAGGTGATTTAAAGCAATCTTAACAGATAT 160  
 DB 589 TCTTAATTTTAT 530  
 QY 161 CAGTTGATTAATAGATTAATTAATATCTGCTTAATGAGAAATATTTGGTTAGCATCA 220  
 DB 529 ATAAATATATAGAAATTTCTTTAGAAATTTTATTAATATAGATTTATTTATTTTAA 470  
 QY 221 GTTGAACAAAGCGGTGAAAATTTCTAAGGAGATATATTTTAAAGATGATGCTGATG 280  
 DB 469 TATTAATATATACATTTTATTAATTAATAGATTAATTAAGTTATTAATTAAGGATA 410  
 QY 281 ATATTCATATCCAGATGATTTGATTAAGCAATTTGTTT-TATGAGGAAATTCATG 339  
 DB 409 TATTAATTTATATATATATATATTTTATATTTTACTTATGATATATATATATTTTAA 350  
 QY 340 GATTTCTACGCAACTTAATAGATTTGATTAAGCAAAAAAGAAATTTAGTATATTAACA 399

Db 349 TACCTTCATATATTTATTTATTTAGTCTAGTAATTTCTATTTAATAGTCTACCT 290  
QY 400 CGAGAAAGTATATAATATCTTAACTAATGATATACGAGATCTTATGATATATCT 459  
Db 289 TTATTTGGATATTTACTACTTAATATTTACCTA-ATATATATTTATTAAGATATCTT 231  
QY 460 ATACTGCCCAACCAAGTGTGCTAAAAAGAAAGTTTCGATAGTTAATGGATAT 519  
Db 230 AATCTATATATTTATCTAAGATATATAATTAATCTTTTATTTATATATTT 171  
QY 520 AGAGATTAGTACCTGTGAAGATATGATTTGCAATAGAGAGCTCGGCTGATTC 579  
Db 170 AATTTATTTATTTATTTAGTAAATTAATTTATTTATTTATTTACATPAATTTTGGATAAT 111  
QY 580 AAATCGGCTTACTCAATAAAGTACTTTTACAGATATGATTAAACA 626  
Db 110 AATATATCATTTAATTAATGTAATTTATTTAATATTTATCTTTAATGA 64

Search completed: March 28, 2004, 07:33:01  
Job time : 119 secs